

Théorie de l'Information et Codage de Source

Exercices de cours corrigés

Olivier Rioul

Janv. 2006

*Polycopié « Information et Codage » par O. Rioul et P. Duhamel
Cours : ENSTA/SIC/ES204 et LMD/SATS/CDS*

I Exercices pour le chapitre 1 : Outils de la théorie de l'information

Débit d'un modem téléphonique.

Sur le réseau NUMERIS le signal de parole est échantillonné à 8 kHz (largeur de bande maximale < 4 kHz par le théorème d'échantillonnage de Nyquist). Chaque échantillon du signal de parole est ensuite quantifié sur 256 niveaux d'amplitude.

1. Calculer le taux de codage R et le débit binaire correspondant en kb/s.

Variable gaussienne.

2. Calculer la moyenne et la variance de la v.a. X de distribution de probabilité :

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

où μ et $\sigma^2 > 0$ sont des paramètres. Commenter.

Formule de Bayes. Traitement réciproque

3. (a) Montrer la formule de Bayes :

$$p(x|y) = \frac{p(y|x)p(x)}{\sum_x p(y|x)p(x)}$$

- (b) Commenter cette formule en considérant le «traitement réciproque» d'entrée Y et de sortie X .

Chaînes de Markov.

4. (a) En s'aidant de la formule $p(x, y) = p(x)p(y|x)$, montrer la même formule conditionnée par z :

$$p(x, y|z) = p(x|z)p(y|x, z)$$

- (b) En déduire que $X \rightarrow Y \rightarrow Z$ forme une chaîne de Markov si et seulement si X et Z sont indépendants sachant y , i.e. :

$$p(x, z|y) = p(x|y)p(z|y)$$

- (c) Montrer que si $X \rightarrow Y \rightarrow Z$ est une chaîne de Markov, alors la chaîne «réciproque» $Z \rightarrow Y \rightarrow X$ l'est aussi.

Positivité de la divergence.

5. (a) Montrer que la fonction logarithme est *strictement concave*.
(b) En déduire l'inégalité de Jensen :

$$\mathbf{E} \log_2 f(X) \leq \log_2 \mathbf{E} f(X)$$

avec égalité si et seulement si la fonction f est constante p.p.

- (c) En considérant $-D(p, q)$ démontrer le résultat fondamental du cours concernant la positivité de la divergence.
(d) Retrouver ce résultat avec l'inégalité $\log_e x \leq x - 1$.

Propriétés de l'entropie.

On considère l'entropie $H(X)$ d'une v.a. discrète X (pouvant prendre un nombre fini M de valeurs).

6. (a) Montrer que $H(X)$ est une «auto-information» $H(X) = I(X, X)$. Interpréter ce résultat.
(b) Montrer que $H(X) \geq 0$ et déterminer le cas d'égalité.
(c) En s'aidant du résultat fondamental sur les divergences (inégalité de Gibbs), montrer que $H(X) \leq \log_2 M$ et déterminer le cas d'égalité.

La connaissance réduit l'incertitude.

7. (a) Montrer et interpréter (dans le cas de v.a. discrète) l'inégalité suivante :

$$H(Y|X) \leq H(Y).$$

- (b) La connaissance réduit-elle toujours l'incertitude ?

Entropie différentielle et entropie absolue.

On considère une v.a. continue X que l'on quantifie uniformément avec un pas de quantification q pour obtenir une v.a. discrète $[X]$.

8. (a) En approximant l'intégrale $H(X) = \int_x p(x) \log_2 \frac{1}{p(x)}$ par une somme de Riemann, établir que

$$H(X) \approx H([X]) - \log_2(1/q)$$

- (b) Interpréter ce résultat lorsque $q \rightarrow 0$ et expliquer le terme «entropie différentielle».

II Exercices pour le chapitre 2 : Application de la théorie de l'information au codage

Démonstration du théorème du traitement de données.

1. (a) A l'aide des deux formules $p(x, y) = p(x)p(y|x)$ et $p(x, y|z) = p(x|z)p(y|x, z)$ (voir exo leçon 1), démontrer la formule :

$$I((X, Y); Z) = I(X, Z) + I(Y, Z|X)$$

- (b) On considère dorénavant une chaîne de Markov $X \rightarrow Y \rightarrow Z$. Développer $I((X, Y); Z)$ de deux manières différentes et en déduire le théorème du traitement de données dans un cas particulier : $I(Y, Z) \geq I(X, Z)$.
- (c) Sachant que $Z \rightarrow Y \rightarrow X$ est également une chaîne de Markov (voir exo leçon 1), montrer que $I(X, Y) \geq I(X, Z)$.
- (d) En déduire l'énoncé général du théorème du traitement de données.

Fonction taux-distorsion : Cas extrêmes.

On considère la fonction taux-distorsion $R(D)$ pour une source sans mémoire.

2. (a) Cas $D = 0$ (codage sans pertes). Etablir la borne de Shannon $R(D = 0)$ dans ce cas. Commenter
- (b) Cas $R = 0$ (pas de transmission). Etablir D à la borne de Shannon $R(D) = 0$ dans ce cas. Commenter.

Entropie d'une source gaussienne.

3. (a) Calculer l'entropie différentielle $H(X)$ lorsque X est une v.a. gaussienne. Commenter.
- (b) Montrer que l'entropie d'une v.a. X de variance σ^2 est maximale lorsque X est gaussienne.

Fonction taux-distorsion : Cas gaussien.

On considère une source gaussienne sans mémoire de moyenne nulle et de variance σ^2 .

4. (a) Montrer que $R(D) = H(X) - \max H(X|Y)$.
- (b) Trouver le maximum de l'entropie conditionnelle sachant que $H(X|Y) = H(X - Y|Y) \leq H(X - Y)$. Justifier que ce maximum peut être effectivement atteint.
- (c) En déduire l'expression cherchée de $R(D)$.

III Exercices pour le chapitre 3 : Codage entropique à longueur variable

Condition du préfixe.

1. (a) Justifier qu'un code vérifiant la condition du préfixe est décodable de manière instantanée.
- (b) Réciproquement, montrer qu'un code instantané vérifie la condition du préfixe.
- (c) En utilisant les résultats du cours sur l'inégalité de Kraft-McMillan, montrer que tout code u.d. peut être remplacé par un code instantané (à préfixe) de même distribution de longueurs et donc de même taux. Commenter.

Démonstration de l'inégalité de Kraft-McMillan.

2. (a) Montrer que pour un code u.d., toute séquence de l bits peut se décomposer d'au plus une façon comme concaténation de mots de codes $c_{i_1} c_{i_2} \dots c_{i_k}$ où $l_{i_1} + l_{i_2} + \dots + l_{i_k} = l$.
- (b) En déduire que le nombre total $N_l(k)$ de concaténations possibles de k mots de codes donnant une séquence codée de longueur totale l bits vérifie l'inégalité : $N_l(k) \leq 2^l$.
- (c) Montrer par ailleurs que

$$\left(\sum_{i=1}^M x^{l_i} \right)^k = \sum_l N_l(k) x^l$$

- (d) Conclure en faisant $x = 1/2$ et $k \rightarrow \infty$.

Construction d'un code instantané.

On se donne une distribution de longueurs $l_1 \leq l_2 \leq \dots \leq l_M$ vérifiant l'inégalité de Kraft-McMillan. A chaque mot de code c_i (à trouver) on associe le nombre

$\bar{c}_i = \overline{0}, c_i \in [0, 1[$ dont les décimales de l'écriture en base 2 est formée des bits de c_i . On note I_i l'intervalle $I_i = [\bar{c}_i; \bar{c}_i + 2^{-l_i}[$.

Par exemple, $c_i = 010$ donne $\bar{c}_i = \overline{0,010} = \frac{1}{4}$. et $I_i = [\overline{0,010}; \overline{0,011}[= [\frac{1}{4}; \frac{3}{8}[$ est l'ensemble des nombres de $[0; 1[$ dont les décimales en base 2 commencent par c_i .

3. (a) Montrer que c_i détermine I_i , et réciproquement.
- (b) Montrer que le code est instantané si et seulement si les I_i sont des intervalles disjoints.
- (c) Interpréter l'inégalité de Kraft-McMillan sur les I_i et en déduire un algorithme de construction du code.
- (d) Préciser cet algorithme sur des exemples pour $l_1 \leq l_2 \leq \dots \leq l_M$, en commençant par $\bar{c}_1 = \overline{0.0\dots 0}$, et en posant $\bar{c}_{i+1} =$ extrémité droite de I_i à chaque étape.
- (e) Que se passe-t-il si les l_i ne vérifient *pas* l'inégalité de Kraft-McMillan ? Donner un exemple.

Algorithme de Huffman. Préliminaires

On considère un code VLC *optimal* pour une source de distribution de probabilité $p_1 \geq p_2 \geq \dots \geq p_M$.

4. (a) Montrer que nécessairement $l_1 \leq l_2 \leq \dots \leq l_M$ (raisonner par l'absurde en supposant $p_i > p_j$ et $l_i > l_j$). Commenter.
- (b) Montrer que nécessairement l'inégalité de Kraft McMillan est une égalité (raisonner par l'absurde en supposant $\sum_i 2^{-l_i} < 1$, et montrer qu'alors on peut remplacer l_M par $l_M - 1$).
- (c) Déduire du raisonnement de la question précédente que $l_{M-1} = l_M$, et qu'on peut toujours se ramener au cas où les deux mots de codes c_{M-1} et c_M ne diffèrent que par le dernier bit.

Algorithme de Huffman.

On considère une source M -aire de distribution de probabilité $p_1 \geq p_2 \geq \dots \geq p_M$. La *réduction de Huffman* consiste à considérer la source $(M-1)$ -aire, dite «réduite», de distribution de probabilité $p_1, p_2, \dots, p_{M-2}, p'_{M-1} = p_{M-1} + p_M$ (on «combine» les deux symboles les moins probables).

On note $\{c_1, \dots, c_{M-1}, c_M\}$ le code optimal cherché (à l'ordre M). D'après ci-dessus, c_{M-1} et c_M ne diffèrent que par le dernier bit; on peut écrire $c_{M-1} = [c'_{M-1} 0]$ et $c_M = [c'_{M-1} 1]$

5. (a) En comparant les taux de codage de la source initiale et de la source réduite après réduction de Huffman, montrer que le code

$$\{c_1, \dots, c_{M-2}, c'_{M-1}\}$$

est optimal pour la source réduite.

- (b) Donner un moyen de construire le code optimal $\{c_1, \dots, c_{M-1}, c_M\}$ à partir de $\{c_1, \dots, c_{M-2}, c'_{M-1}\}$.
- (c) Par réductions de Huffman successives jusqu'au cas $M = 2$ (où le code $\{0,1\}$ est optimal), obtenir un algorithme de construction du code $\{c_1, \dots, c_{M-1}, c_M\}$.

Indication : Il faut réordonner à chaque étape les probabilités après chaque réduction de Huffman.

IV Exercices pour le chapitre 4 : Quantification scalaire.

Caractéristique Débit-Distorsion en haute résolution

On effectue une quantification scalaire uniforme haute résolution d'une source quelconque de variance σ^2 et de densité de probabilité à support borné dans $[-A, A]$.

1. (a) Calculer la distorsion en fonction du taux de quantification. On écrira le résultat en fonction du facteur $\gamma = \frac{A}{\sigma}$.
- (b) Expliquer le terme «6 dB par bit» pour qualifier la caractéristique débit/distorsion.

Minima locaux de l'algorithme de Lloyd-Max.

On considère un signal aléatoire X de densité de probabilité :

$$p(x) = \begin{cases} 1 & \text{pour } 1 \leq x \leq 1.5 \\ 2 & \text{pour } -1.25 \leq x \leq -1 \\ 0 & \text{sinon.} \end{cases}$$

Il est quantifié scalairement sur 3 niveaux. On considère les deux solutions suivantes caractérisées par la position des centroïdes :

$$\left\{-1 - \frac{3}{16}, -1 - \frac{1}{16}, 1.25\right\} \text{ et } \left\{-1 - \frac{1}{8}, 1 + \frac{1}{8}, 1 + \frac{3}{8}\right\}$$

2. (a) Donner, pour chacune de ces solutions, les cellules de quantification optimales.
- (b) Vérifier que ces deux situations vérifient les conditions de convergence (point stationnaire) de l'algorithme de Lloyd-Max.
- (c) Calculer la contribution à la distorsion quadratique d'une cellule du type $[-q, q]$ correspondant à une amplitude A de la densité de probabilité.

- (d) En déduire les valeurs des distorsions totales dans les deux cas considérés. Quelle est la solution la meilleure ?
- (e) Qu'en déduire sur l'algorithme de Lloyd-Max ? La fonction $\log p(x)$ est-elle concave ?

Densité des cellules de quantification

- 3. Justifier que $\lambda(x) = \frac{1}{Mq(x)}$ représente la «densité» des cellules en évaluant l'intégrale $\int_I \lambda(x) dx$ prise sur un intervalle I .

Optimisation de la formule de Bennett

- 4. (a) Dans la formule de Bennett, quelle est la densité des cellules $\lambda(x) = \frac{1}{Mq(x)}$ qui minimise la distorsion ? Indication : Ecrire le lagrangien correspondant.
- (b) Calculer la distorsion minimale correspondante.
- (c) Appliquer ces résultats à la source gaussienne.

Réalisation de la quantification scalaire non uniforme par non-linéarités.

On réalise une quantification scalaire non uniforme d'une source U de la manière suivante : On transforme d'abord la source $X = f(U)$ à l'aide d'une fonction non linéaire f . On applique ensuite une quantification scalaire uniforme à X qui fournit Y , et on applique enfin la non-linéarité inverse $V = f^{-1}(Y)$.

- 5. (a) Faire un dessin.
- (b) Sous les hypothèses de haute résolution, déterminer la densité des cellules $\lambda(u) = \frac{1}{Mq(u)}$ en fonction de la non-linéarité f .
- (c) En déduire la non-linéarité optimale qui rend la distorsion quadratique minimale, en fonction de la densité de probabilité de la source $p(u)$. Indication : Utiliser l'exercice précédent.

Quantification vectorielle

En quantification vectorielle en dimension n , on attribue à chaque vecteur de source $\underline{X} = (X_1, X_2, \dots, X_n)$ une étiquette binaire correspondant à un centroïde \underline{Y} en dimension n .

- 6. (a) Reprendre, dans le cas vectoriel, les conditions du plus proche voisin et du centroïde vues en cours.
- (b) En déduire l'algorithme de Lloyd-Max dans ce cas.
- (c) Montrer que l'algorithme de Lloyd-Max peut converger vers un minimum local, même si $\log p(\underline{x})$ est concave.

Indication : On considérera pour cela une source uniforme X dans l'intervalle $[-1, 1]$ quantifiée en dimension 2 sur 1 bit par échantillon (c'est à dire 4 centroïdes en 2 dimensions) et les deux situations suivantes :

$$\text{centroïdes } \underline{y} = (\pm \frac{1}{2}, \pm \frac{1}{2})$$

et

$$\text{centroïdes } \underline{y} = (\pm \frac{1}{4}, 0) \text{ et } (\pm \frac{3}{4}, 0).$$

V Exercices pour le chapitre 5 : Codage par transformée

Gain de Codage

1. (a) Donner l'inégalité de concavité du logarithme.
- (b) En déduire que le gain de codage par transformée est toujours ≥ 1 . Quel est le cas d'égalité ?
- (c) Donner un exemple de source pour laquelle le gain de codage est toujours = 1.

Inégalité de Hadamard

Soit \underline{X} un vecteur aléatoire, de composantes X_i de variance σ_i^2 . On définit la matrice d'autocovariance :

$$\underline{\underline{R}} = \mathbf{E}(\underline{X} \cdot \underline{X}^t).$$

2. (a) Donner les coefficients $r_{i,j}$ de la matrice $\underline{\underline{R}}$.
- (b) Quelle est la forme particulière de $\underline{\underline{R}}$ pour des composantes X_i décorréélées ?
- (c) Les composantes réduites sont $X'_i = \frac{X_i}{\sigma_i}$. Donner la matrice d'autocovariance $\underline{\underline{R'}}$ des $\underline{X'_i}$ en fonction de $\underline{\underline{R}}$.
- (d) En déduire $\det \underline{\underline{R}}$ en fonction de $\det \underline{\underline{R'}}$ et des σ_i^2 .
- (e) Montrer que $\det \underline{\underline{R'}} \leq 1$ (raisonner sur les valeurs propres de $\underline{\underline{R'}}$).
- (f) En déduire l'inégalité de Hadamard :

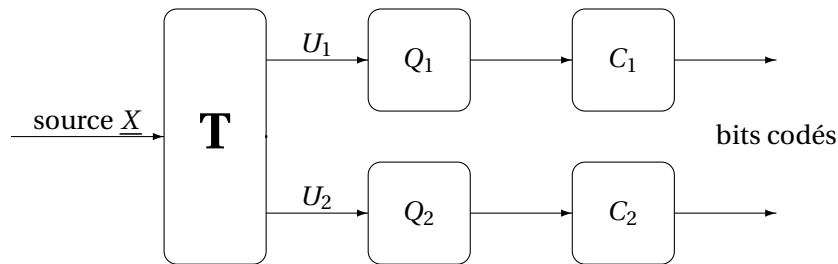
$$\det \underline{\underline{R}} \leq \prod_i \sigma_i^2$$

Quel est le cas d'égalité ?

Transformée de Karhunen-Loève

3. (a) En considérant le gain de codage, montrer que la transformée optimale est celle qui minimise le produit des variances en sortie de la transformée.
- (b) A l'aide de l'inégalité de Hadamard, montrer que la transformée optimale est celle qui décorrèle la sortie.
- (c) Comment obtenir la transformée optimale à partir de $R_{\underline{X}\underline{X}} = \mathbf{E}(\underline{X}\underline{X}^t)$? Cette transformée s'appelle la transformée de Karhunen-Loève.
- (d) Expliciter le gain de codage optimal.
- (e) Pour quel type de source le codage par transformée s'avère-t-il inutile? Est-il pour autant nuisible?

Codage par transformée pour $n = 2$



On considère un schéma de codage de source par transformée orthogonale dont la partie «codage» est représentée dans la figure. Chaque bloc de source $\underline{X} = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$ est transformée en deux échantillons $\begin{pmatrix} U_1 \\ U_2 \end{pmatrix} = \underline{U}$ avant d'être quantifié et codé. On a donc :

$$\underline{U} = \mathbf{T} \cdot \underline{X}$$

où \mathbf{T} est la matrice de la transformée orthogonale ($\mathbf{T}\mathbf{T}^t = \mathbf{I}$).

Les quantificateurs sont scalaires uniformes et les codeurs entropiques sont des codeurs de Huffman. Pour chaque branche i ($i = 1, 2$), la distorsion quadratique moyenne dûe au quantificateur Q_i est notée D_i et le taux binaire moyen après codage de Huffman C_i est noté R_i .

On modélise les signaux en sortie de transformée par des sources laplaciennes de variances σ_1^2 et σ_2^2 . On admet la relation :

$$D_i = c\sigma_i^2 2^{-2R_i} \quad \text{pour } i = 1, 2$$

où c est une constante.

4. (a) Justifier, d'après le cours, que les distorsion quadratique moyenne globale D et taux global R sont donnés par :

$$D = \frac{D_1 + D_2}{2} \quad \text{et} \quad R = \frac{R_1 + R_2}{2}$$

- (b) Justifier (sans calcul), d'après le cours, qu'après optimisation des taux R_1 et R_2 la distorsion minimisée est donné par la formule :

$$D^* = c\sigma_1\sigma_22^{-2R}$$

La matrice d'autocorrélation du signal d'entrée \underline{X} est donnée par :

$$\mathbf{R} = \mathbf{E}(\underline{X}\underline{X}^t) = \sigma_x^2 \cdot \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$$

où ρ est un coefficient de corrélation ($-1 < \rho < 1$).

- (c) Calculer la matrice d'autocorrélation de \underline{U} en fonction de \mathbf{R} et \mathbf{T} .
(d) En déduire la matrice d'autocorrélation de \underline{U} lorsque

$$\mathbf{T} = \beta \cdot \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

(On déterminera d'abord la valeur de β pour que \mathbf{T} soit orthogonale).

- (e) Justifier que la transformée optimale est celle donnée à la question précédente. Sous quel nom est-elle connue ?
(f) Donner l'expression de la distorsion D^* pour cette transformée en fonction de c , σ_x^2 , ρ et R .
(g) Qu'observe-t-on si ρ augmente ? Commenter.

Solutions

I-1) $M = 256$ niveaux en dimension $n = 1$ d'où $R = \frac{\log_2 M}{n} = 8$ bits/éch. Fréquence d'échantillonnage $\frac{1}{T} = 2 \times 4 = 8$ kHz d'où un débit $= \frac{R}{T} = 64$ kb/s.

I-2) Moyenne $\mathbf{E}(X) = \mathbf{E}(X - \mu) + \mu$. Le premier terme est nul (intégrale d'une fonction impaire), d'où $\mathbf{E}(X) = \mu$.

Variance $\mathbf{E}((X - \mu)^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \int x^2 e^{-\frac{x^2}{2\sigma^2}}$ par changement de variable. Une intégration par parties donne $\mathbf{E}((X - \mu)^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \int \sigma^2 e^{-\frac{x^2}{2\sigma^2}} = \sigma^2$

On retrouve bien les paramètres de moyenne μ et de variance σ^2 .

I-3a) $p(x|y) = \frac{p(x,y)}{p(y)}$ avec $p(x,y) = p(y|x)p(x)$ et $p(y) = \int p(x,y) = \int p(y|x)p(x)$, d'où la formule.

I-3b) Le traitement direct d'entrée X et de sortie Y est caractérisé par $p(y|x)$. Le traitement réciproque (d'entrée Y et de sortie X) est caractérisé par $p(x|y)$ donné par la formule ci-dessus : ce traitement dépend non seulement du traitement direct, mais aussi de la distribution de probabilité de X (exemple : racine carrée probabiliste).

I-4a) $p(x,y|z) = \frac{p(x,y,z)}{p(z)} = \frac{p(x,z)}{p(z)} p(y|x,z) = p(x|z)p(y|x,z)$. C'est la même formule que $p(x,y) = p(x)p(y|x)$, mais conditionnée par z .

I-4b) La condition de Markov est $p(z|x,y) = p(z|y)$. Or $p(z|x,y)p(x|y) = p(x,z|y)$ d'après la question précédente. D'où la condition équivalente de l'énoncé en remplaçant $p(z|x,y)$.

I-4c) X et Z sont indépendants sachant Y si et seulement si Z et X sont indépendants sachant Y , d'où le résultat.

I-5a) La dérivée seconde de $\log x$ est $-1/x^2 < 0$.

I-5b) C'est une inégalité de concavité stricte (Jensen) : la moyenne d'un log est inférieure au log de la moyenne, avec égalité si et seulement si on moyenne sur une constante.

I-5c) $-D(p,q) = \mathbf{E} \log \frac{q(X)}{p(X)} \leq \log \mathbf{E} \frac{q(X)}{p(X)} = \log 1 = 0$ avec égalité ssi $q(x) \propto p(x)$ p.p., c'est à dire $p = q$ (même normalisation).

I-5d) $\mathbf{E} \log \frac{q(X)}{p(X)} \leq \mathbf{E} \frac{q(X)}{p(X)} - 1 = \int q(x) - 1 = 0$

I-6a) $I(X,X) = \sum_x p(x,x) \log_2 \frac{p(x,x)}{p(x)p(x)}$, or $p(x,x) = \mathcal{P}rob\{X = x \text{ et } X = x\} = p(x)$, d'où après simplification : $I(X,X) = \sum_x p(x) \log_2 \frac{1}{p(x)} = H(X)$. L'entropie est la quantité d'information qu'apporte, en moyenne, une réalisation de X sur la connaissance de X elle-même.

I-6b) On sait que l'information mutuelle est positive ou nulle, donc $H(X) = I(X, X) \geq 0$. Il y a égalité $H(X) = I(X, X) = 0$ si et seulement si X est indépendant de lui-même, c'est à dire $p(x, x) = p(x)p(x)$, ou $p(x) = 0$ ou 1 , ou encore : X est une v.a. déterministe.

I-6c) L'inégalité de Gibbs s'écrit $\sum_x p(x) \log_2 \frac{1}{p(x)} \leq \sum_x p(x) \log_2 \frac{1}{q(x)}$, d'où en prenant $q(x) = \frac{1}{M}$: $H(X) \leq \log_2 M$. Il ya égalité ssi $p(x) = q(x) = \frac{1}{M}$ pour tout x , c'est à dire que X est uniforme.

I-7a) $H(Y) - H(Y|X) = I(X, Y) \geq 0$ Connaître X réduit l'incertitude moyenne sur Y .

I-7b) Toujours en moyenne, d'après la question précédente. Mais on peut avoir $H(Y|X = x) \geq H(Y)$ pour un événement $X = x$ particulier.

I-8a) $\int p(x) \log_2 \frac{1}{p(x)} dx \approx \sum_i p(x_i) \log_2 \frac{1}{p(x_i)} q = H([X]) - \log_2(1/q)$.

I-8b) Quand $q \rightarrow 0$, $\log_2 \frac{1}{q} \rightarrow \infty$ et donc $H([X]) \rightarrow \infty$ si l'entropie différentielle est finie. L'entropie différentielle est donc une *différence* de deux entropies ($H([x])$ et $\log_2 \frac{1}{q}$) chacune tendant vers l'infini quand la précision augmente.

II-1a) On divise membre à membre les deux formules :

$$\frac{p(x, y|z)}{p(x, y)} = \frac{p(x|z)}{p(x)} \frac{p(y|x, z)}{p(y|x)}$$

et on prend la moyenne du logarithme, on obtient bien $I((X, Y); Z) = I(X, Z) + I(Y, Z|X)$.

II-1b) D'une part $I((X, Y); Z) = I(X, Z) + I(Y, Z|X)$ et d'autre part, en échangeant les rôles de X et Y : $I((X, Y); Z) = I(Y, Z) + I(X, Z|Y)$ avec $I(X, Z|Y) = 0$ par la condition de Markov. D'où en identifiant $I(Y, Z) = I(X, Z) + I(Y, Z|X) \geq I(X, Z)$.

II-1c) Echanger les rôles de X et Z .

II-1d) Soit une chaîne de Markov quelconque $\rightarrow X \rightarrow \dots \rightarrow U \rightarrow \dots \rightarrow V \rightarrow \dots \rightarrow Y$. Alors $I(X, Y) \leq I(U, Y)$ (considérer la sous-chaîne $X \rightarrow U \rightarrow Y$) puis $I(U, Y) \leq I(U, V)$ (considérer la sous-chaîne $U \rightarrow V \rightarrow Y$). On obtient bien l'inégalité générale $I(X, Y) \leq I(U, V)$.

II-2a) $D = 0$ revient à dire $X = Y$, d'où $I(X, Y) = I(X, X)$ qui est l'entropie $H(X)$ dans le cas discret (et ∞ dans le cas continu). On trouve donc $R(D = 0)$ égal à l'entropie de la source (codage sans pertes).

II-2b) $R(D) = 0$ implique qu'au minimum sur $p(y|x)$, $I(X, Y) = 0$: X et Y sont indépendants, $p(y|x) = p(y)$, avec $D \geq \mathbf{E}((X - Y)^2) = \int_y p(y) \mathbf{E}_X((X - y)^2)$ Le minimum de l'espérance sur X s'obtient pour $y = \mathbf{E}(X)$ (auquel cas ce minimum est égal à la variance de X). En prenant Y constant égal à $\mathbf{E}(X)$, on obtient donc $D = \sigma^2$ comme valeur minimale de D . Ainsi $R(D) = 0 \iff D \geq \sigma^2$

II-3a) $H(X) = \mathbf{E} \log_2 \frac{1}{p(X)} = \mathbf{E} \log_2 \sqrt{2\pi\sigma^2} \exp \frac{(X-\mu)^2}{2\sigma^2} = \frac{1}{2} \log_2(2\pi\sigma^2) + (\log_2 e) \mathbf{E} \frac{(X-\mu)^2}{2\sigma^2} = \frac{1}{2} \log_2(2\pi\sigma^2) + \frac{1}{2} \log_2 e = \left[\frac{1}{2} \log_2(2\pi e\sigma^2) \right]$. Cette valeur est positive si $\sigma^2 \geq \frac{1}{2\pi e}$, négative sinon. L'entropie différentielle n'est pas nécessairement positive.

II-3b) L'inégalité de Gibbs s'écrit $\int p(x) \log_2 \frac{1}{p(x)} dx \leq \int p(x) \log_2 \frac{1}{q(x)} dx$, d'où en prenant $q(x)$ = gaussienne, $H(X) \leq \frac{1}{2} \log_2(2\pi e\sigma^2)$ (même calcul que pour la question précédente). Il y a égalité si et seulement si $p = q$, c'est à dire si X est gaussien.

II-4a) $R(D) = \min I(X, Y) = \min H(X) - H(X|Y) = H(X) - \max H(X|Y)$ car $H(X) = \frac{1}{2} \log_2(2\pi e\sigma^2)$ est constant.

II-4b) $H(X|Y) = H(X - Y|Y)$ par changement de variable, puis $H(X - Y|Y) \leq H(X - Y)$ (la connaissance réduit l'incertitude, avec égalité si et seulement si l'erreur $X - Y$ est indépendante de Y), puis $H(X - Y) \leq \frac{1}{2} \log_2(2\pi eD)$ avec égalité si et seulement si l'erreur $X - Y$ est gaussienne.

En résumé, le maximum de l'entropie conditionnelle est effectivement atteint si $X - Y$ peut être gaussien de variance D , indépendant de Y . C'est possible lorsque $D \leq \sigma^2$, car

$$X = (X - Y) + Y$$

donne alors une somme de v.a. gaussiennes indépendantes qui est bien une gaussienne. Pour que la variance de X soit σ^2 , il faut que celle de Y soit $= \sigma^2 - D \geq 0$. Par ailleurs si $D \geq \sigma^2$, on sait que $R(D) = 0$ (cf. ci-dessus).

II-4c) On trouve $R(D) = \frac{1}{2} \log_2(2\pi e\sigma^2) - \frac{1}{2} \log_2(2\pi eD)$ pour $D \leq \sigma^2$, d'où

$$R(D) = \begin{cases} \frac{1}{2} \log_2 \frac{\sigma^2}{D} & \text{si } D \leq \sigma^2 \\ 0 & \text{sinon.} \end{cases}$$

III-1a) Dès que le décodeur a lu l_i bits d'un mot de code c_i , aucun autre mot de code que c_i ne peut être décodé, et c_i est décodé instantanément.

III-1b) Puisque le décodeur décode instantanément tout mot de code c_i , aucun autre mot de code ne peut commencer par les l_i bits de c_i .

III-1c) Tout code u.d. admet une distribution de longueur vérifiant l'inégalité de Kraft-McMillan. Puisque cette inégalité est satisfaite, il existe (construction de Kraft) un code instantané de même distribution de longueurs et donc de même taux. On peut donc toujours se restreindre à la recherche de codes instantanés (si le critère porte sur le taux uniquement).

III-2a) C'est la définition même du décodage instantané.

III-2b) Le nombre total de suites de l bits est 2^l , et d'après la question précédente, chaque suite de ce type peut se décomposer d'au plus une façon comme concaténation de k mots de codes. D'où $N_l(k) \leq 2^l$.

III-2c)

$$\begin{aligned} \left(\sum_{i=1}^M x^{l_i}\right)^k &= \sum_{i_1} \sum_{i_2} \dots \sum_{i_k} x^{l_{i_1} + l_{i_2} + \dots + l_{i_k}} \\ &= \sum_l N_l(k) x^l \quad \text{en regroupant suivant } l \end{aligned}$$

III-2d) On trouve $(\sum_i 2^{-l_i})^k \leq \sum_l N_l(k) 2^{-l}$, or cette dernière somme est bornée par $k \cdot (l_{\max} - 1)$ où l_{\max} est la longueur maximale d'un mot de code. Ainsi quand $k \rightarrow \infty$:

$$\sum_i 2^{l_i} \leq (ck)^{1/k} \rightarrow 1$$

d'où l'inégalité cherchée.

III-3a) c_i détermine \bar{c}_i et l_i donc $I_i = [\bar{c}_i; \bar{c}_i + 2^{-l_i}[$. Réciproquement, cette intervalle détermine l_i et les l_i première décimales de \bar{c}_i s'identifient à c_i .

III-3b) La condition du préfixe revient à dire qu'aucun \bar{c}_i ne se trouve dans un I_j , $j \neq i$, autrement I_i et I_j sont disjoints $\forall i \neq j$.

III-3c) La longueur de l'intervalle I_i est 2^{-l_i} , et l'inégalité de Kraft-McMillan signifie que la somme totale des longueurs des intervalles I_i est inférieure ou égale à 1 (longueur du segment $[0, 1[$). Il est donc possible choisir les I_i de telle sorte qu'ils soient tous inclus dans le segment $[0, 1[$, ce qui fournit un code instantané.

III-3d) On met bout à bout les intervalles par ordre décroissant de longueurs dans le segment $[0, 1[$. Cela revient à commencer par le mot de code 0, puis à ajouter 1 (en base 2) à chaque étape, en remplissant si nécessaire à droite par des zéros de façon à trouver la bonne longueur. Voici un exemple :

1	0
2	10
4	1100
4	1101
4	1110
5	11110
6	111110

III-3e) On obtient une impossibilité (on déborde de l'intervalle $[0, 1]$). Exemple :

1	0
2	10
3	110
4	1110
4	1111
4	impossible

III-4a) Si on avait disons $p_i > p_j$ et $l_i > l_j$, il suffirait d'échanger c_i et c_j pour obtenir un meilleur code (car $p_i l_i + p_j l_j > p_i l_j + p_j l_i$). C'est absurde.

III-4b) Si $\sum_i 2^{-l_i} = N 2^{-l_M} < 1$, alors l'entier N vérifie $N < 2^{l_M}$ c'est à dire $N \leq 2^{l_M} - 1$ et donc $\sum_i 2^{-l_i} \leq 1 - 2^{-l_M}$, ou encore $\sum_i 2^{-l_i} + 2^{-l_M} \leq 1$. On peut donc remplacer l_M par $l_M - 1$ en conservant l'inégalité de Kraft-McMillan. On obtiendrait ainsi un meilleur code (le taux est réduit), ce qui est absurde.

III-4c) Puisque $\sum_i 2^{-l_i} = N 2^{-l_M} = 1$, on a $N = 2^{l_M}$. Or N est une somme de puissances de 2 :

$$N = 2^{l_M - l_1} + \dots + 2^{l_M - l_{M-1}} + 1$$

qui est un nombre pair. On doit donc avoir $l_{M-1} = l_M$. Avec la construction de l'exercice précédent, les deux derniers mots de codes seront de la forme $c_{M-1} = [c'_{M-1} 0]$ et $c_M = [c'_{M-1} 1]$.

III-5a) A l'ordre M :

$$R_M = \sum_i p_i l_i = p_1 l_1 + \dots + p_{M-2} l_{M-2} + l_M p'_{M-1}$$

et pour le code réduit :

$$R_{M-1} = p_1 l_1 + \dots + p_{M-2} l_{M-2} + (l_M - 1) p'_{M-1}$$

Ces deux taux ne diffèrent que de la constante p'_{M-1} . Ainsi, R_M étant minimal, R_{M-1} l'est aussi pour la source réduite.

III-5b) D'après l'exercice précédent il suffit de choisir $c_{M-1} = [c'_{M-1} 0]$ et $c_M = [c'_{M-1} 1]$.

III-5c) On réduit successivement en réordonnant à chaque étape les probabilités, jusqu'à obtenir $M = 2$, auquel cas le code optimal est nécessairement $\{0, 1\}$. Puis en tenant compte des ré-ordonnements faits on attribue successivement les mots de code comme à la question précédente. Exemple :

Descente. La colonne de gauche donne la distribution de probabilité initiale ($M = 8$). Chaque réduction crée une nouvelle colonne à droite (après réarrangement par ordre décroissant) :

0.25	0.25	0.25	0.25	0.3	0.45	0.55
0.25	0.25	0.25	0.25	0.25	0.3	0.45
0.2	0.2	0.2	0.2	0.25	0.25	
0.14	0.14	0.14	0.16	0.2		
0.1	0.1	0.1	0.14			
0.04	0.04	0.06				
0.01	0.02					
0.01						

Remontée. De la droite vers la gauche, on construit de proche en proche le code optimal en tenant compte des permutations faites (on peut utiliser le même tableau si on le fait à la main) :

01	01	01	01	00	1	0
10	10	10	10	01	00	1
11	11	11	11	10	01	
001	001	001	000	11		
0000	0000	0000	001			
00010	00010	0001				
000110	00011					
000111						

IV-1a) $D = \frac{q^2}{12}$ avec $q = \frac{2A}{M}$ et $M = 2^R$, d'où

$$D = \frac{1}{12} \left(\frac{2A}{2^R} \right)^2 = \boxed{\frac{\gamma^2}{3} \sigma^2 2^{-2R}}$$

IV-1b) Si on attribue un bit supplémentaire au quantificateur ($R \rightarrow R + 1$), la distortion est divisée par 4 (-6dB) et donc le rapport signal à bruit augmente de 6dB par bit.

IV-2a) Les frontières des cellules optimales sont à mi-distance des centroïdes. Solution 1 : $-1 - \frac{2}{16}$ et 0 (ou tout autre frontière dans la zone intermédiaire pour la densité de probabilité est nulle). Solution 2 : 0 (même remarque) et $1 + \frac{1}{4}$.

IV-2b) La condition du plus proche voisin est satisfaite par la question précédente. Celle du centroïde découle du fait que la distribution de probabilité est constante dans chaque des deux intervalles $1 \leq x \leq 1.5$ et $-1.25 \leq x \leq -1$.

IV-2c) $\int_{-q}^q Ax^2 dx = A \frac{2q^3}{3}$

IV-2d) Solution 1 : $2 \times 2 \frac{2(1/16)^3}{3} + \frac{2(1/4)^3}{3}$ Solution 2 : $2 \frac{2(1/8)^3}{3} + 2 \times \frac{2(1/8)^3}{3}$ La solution 2 est clairement meilleure.

IV-2e) L'algorithme de Lloyd-Max peut converger vers deux solutions distinctes. D'après le cours, ce ne serait pas le cas si $\log p(x)$ était concave.

IV-3)

$$\begin{aligned}\int_I \lambda(x) dx &= \sum_{i; R_i \subset I} \int_{R_i} \frac{1}{M q_i} dx = \sum_{i; R_i \subset I} q_i \frac{1}{M q_i} \\ &= \frac{1}{M} \sum_{i; R_i \subset I} 1 = \frac{|\{i; R_i \subset I\}|}{M}\end{aligned}$$

C'est la proportion du nombre de cellules contenues dans I (rapporté au nombre total = M de cellules).

IV-4a) La distorsion donnée par la formule de Bennett est minimale si $\int \frac{p(x)}{\lambda^2(x)} dx$ est minimum, pour une densité de cellules vérifiant la contrainte $\int \lambda(x) dx = 1$. Le lagrangien correspondant est

$$\mathcal{L} = \int \frac{p(x)}{\lambda^2(x)} dx + \mu \int \lambda(x) dx$$

où μ est le multiplicateur de Lagrange associé à la contrainte. En dérivant par rapport à $\lambda(x)$, on obtient

$$\frac{\partial \mathcal{L}}{\partial \lambda(x)} = -2 \frac{p(x)}{\lambda^3(x)} + \mu = 0$$

et donc à l'optimum $\lambda(x) \propto p(x)^{1/3}$. La constante de proportionnalité est déterminée par la contrainte de normalisation. On doit donc avoir

$$\lambda(x) = \frac{p(x)^{1/3}}{\int p(x)^{1/3} dx}$$

IV-4b) A l'optimum

$$\int \frac{p(x)}{\lambda^2(x)} dx = \left(\int p(x)^{1/3} dx \right)^3$$

et la formule de Bennett devient

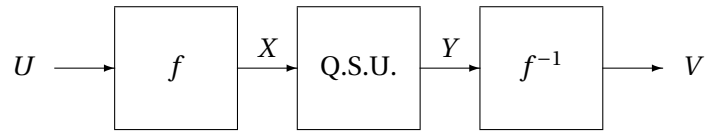
$$D = \frac{1}{12} \left(\int p(x)^{1/3} dx \right)^3 2^{-2R}$$

IV-4c) Pour une source gaussienne $p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}}$ on trouve

$$\left(\int p(x)^{1/3} dx \right)^3 = \frac{1}{\sqrt{2\pi\sigma^2}} \left(\int e^{-\frac{x^2}{2 \times 3 \sigma^2}} dx \right)^3 = \frac{\sqrt{6\pi\sigma^2}^3}{\sqrt{2\pi\sigma^2}}$$

d'où $D = \frac{\pi\sqrt{3}}{2} \sigma^2 2^{-2R}$.

IV-5a) Voilà



IV-5b) $\lambda(u) = \frac{1/M}{q(u)}$ représente le rapport des pas de quantification uniforme (sur X) et non-uniforme (sur U). En haute résolution, on peut approximer ceci par la pente infinitésimale de la non linéarité = $f'(u)$

IV-5c) On sait qu'à l'optimum $\lambda(u) \propto p(u)^{1/3}$ d'où $f(u) = \int^u p(u)^{1/3} du + c$

IV-6a) Plus proche voisin :

$$R_i = \left\{ \underline{x} \mid \|\underline{x} - \underline{y}_i\| \leq \|\underline{x} - \underline{y}_j\| \forall j \right\}$$

Centroïde :

$$y_i = \frac{\int_{R_i} \underline{x} p(\underline{x}) d\underline{x}}{\int_{R_i} p(\underline{x}) d\underline{x}}$$

IV-6b) Comme dans le cas scalaire avec les deux conditions ci-dessus.

IV-6c) Ici $p(\underline{x})$ est constante dans un carré, donc log-concave. Il est immédiat de voir qu'en imposant la condition du plus proche voisin, les deux solutions vérifient la condition du centroïde. On trouve que la première solution est meilleure. L'algorithme de Lloyd-Max peut donc converger vers deux solutions distinctes.

V-1a) $\mathbf{E} \log X \leq \log \mathbf{E} X$ (Jensen), c'est à dire en discret :

$$\sum_i p_i \log x_i \leq \log \sum_i p_i x_i$$

avec $p_i \geq 0$ et $\sum_i p_i = 1$. Comme le logarithme est strictement concave, il y a égalité si et seulement si $X = x_i = \text{constante}$.

V-1b) Prendre $p_i = \frac{1}{n}$ dans la question précédente. En passant à l'exponentielle :

$$\sqrt[n]{\prod_i x_i} \leq \frac{1}{n} \sum_i x_i$$

(moyenne géométrique \leq moyenne arithmétique). Pour $x_i = \sigma_i^2$, il vient $G_{TC} \geq 1$. Le cas d'égalité est que les σ_i^2 sont tous égaux.

V-1c) Une source sans mémoire (bruit blanc) a pour matrice d'autocorrélation = $\sigma^2 \mathbf{I}$. Après transformée orthogonale \mathbf{T} , cette source reste blanche puisque de matrice d'autocorrélation = $\mathbf{T} \sigma^2 \mathbf{T}^t = \sigma^2 \mathbf{I}$. Elle vérifie donc $\sigma_i^2 = \sigma^2$ et $G_{TC} = 1$. On ne peut pas améliorer le gain de codage d'une source blanche par transformée.

V-2a) $r_{i,j} = \mathbf{E}(X_i X_j)$.

V-2b) $r_{i,j} = \mathbf{E}(X_i X_j) = 0$ pour $i \neq j$ et $r_{i,i} = \mathbf{E} X_i^2 = \sigma_i^2$, d'où $\underline{\underline{R}}$ est diagonale.

V-2c) Posons $\underline{\underline{\Sigma}}$ la matrice diagonale d'éléments diagonaux σ_i . On a

$$\underline{\underline{R}}' = \mathbf{E}(\underline{\underline{\Sigma}}^{-1} \underline{\underline{X}} \cdot (\underline{\underline{\Sigma}}^{-1} \underline{\underline{X}})^t) = \underline{\underline{\Sigma}}^{-1} \underline{\underline{R}} \underline{\underline{\Sigma}}^{-1}$$

V-2d) $\det \underline{\underline{R}}' = \frac{\det \underline{\underline{R}}}{(\det \underline{\underline{\Sigma}})^2}$ d'où $\det \underline{\underline{R}} = \det \underline{\underline{R}}' \cdot \prod_i \sigma_i^2$.

V-2e) Ce déterminant est le produit des valeurs propres $\prod_i \lambda_i$ de $\underline{\underline{R}}'$. Or la somme $\sum_i \lambda_i$ de ces valeurs propres est la trace, égale à la somme des éléments diagonaux qui valent tous 1. Donc $\sum_i \lambda_i = n$ et

$$\sqrt[n]{\prod_i \lambda_i} \leq \frac{1}{n} \sum_i \lambda_i = 1$$

c'est à dire $\det \underline{\underline{R}}' \leq 1$.

V-2f) D'après les questions précédentes $\det \underline{\underline{R}} = \det \underline{\underline{R}}' \cdot \prod_i \sigma_i^2 \leq \prod_i \sigma_i^2$. Il y a égalité si les valeurs propres de $\underline{\underline{R}}'$ sont toutes égales, c'est à dire si $\underline{\underline{R}}' = \underline{\underline{I}}$ (matrice identité). Cela revient à dire qu'il y a égalité si et seulement si les composantes X_i sont *décorrélées*.

V-3a) Le gain de codage s'écrit $G_{\text{TC}} = \frac{\frac{1}{n} \sum_i \sigma_i^2}{\sqrt[n]{\prod_i \sigma_i^2}} = \frac{\sigma^2}{\sqrt[n]{\prod_i \sigma_i^2}}$ où σ^2 est la variance de la source. Ce gain est maximum lorsque $\prod_i \sigma_i^2$ est minimum.

V-3b) $\prod_i \sigma_i^2$ est minimum lorsque l'inégalité de Hadamard est une égalité : $\prod_i \sigma_i^2 = \det \underline{\underline{R}}_{\underline{\underline{U}}, \underline{\underline{U}}}$, ce qui revient à dire que les échantillons U_i en sortie de la transformée sont *décorrélés*.

V-3c) On $\underline{\underline{R}}_{\underline{\underline{U}}, \underline{\underline{U}}} = \mathbf{E} \underline{\underline{U}} \underline{\underline{U}}^t$ avec $\underline{\underline{U}} = \underline{\underline{T}} \underline{\underline{X}}$, d'où $\underline{\underline{R}}_{\underline{\underline{U}}, \underline{\underline{U}}} = \underline{\underline{T}} \underline{\underline{R}}_{\underline{\underline{X}}, \underline{\underline{X}}} \underline{\underline{T}}^{-1}$. Cette matrice est diagonale si et seulement si $\underline{\underline{T}}$ diagonalise $\underline{\underline{R}}_{\underline{\underline{X}}, \underline{\underline{X}}}$. Plus précisément, les lignes de $\underline{\underline{T}}$ forme une base orthonormée de vecteurs propres de $\underline{\underline{R}}_{\underline{\underline{X}}, \underline{\underline{X}}}$.

V-3d) A l'optimum, les σ_i^2 sont les valeurs propres de $\underline{\underline{R}}_{\underline{\underline{X}}, \underline{\underline{X}}}$. Donc

$$G_{\text{TC}} = \frac{\frac{1}{n} \sum_i \sigma_i^2}{\sqrt[n]{\prod_i \sigma_i^2}} = \frac{\frac{1}{n} \text{Tr}(\underline{\underline{R}}_{\underline{\underline{X}}, \underline{\underline{X}}})}{\sqrt[n]{\det \underline{\underline{R}}_{\underline{\underline{X}}, \underline{\underline{X}}}}.$$

V-3e) La transformée optimale *décorrèle* la source, le codage par transformée est donc inutile lorsque les échantillons de source sont déjà *décorrélés*. Dans ce cas on a vu que $G_{\text{TC}} = 1$ pour toute transformée orthogonale. La codage par transformée n'est donc pas nuisible.

V-4a) Voir le cours ($D = \frac{D_1+D_2}{2}$ parce que la transformée est orthogonale).

V-4b) Voir le cours : $D = D_1 = D_2 = \sqrt{D_1 D_2} = c\sqrt{\sigma_1^2 \sigma_2^2} 2^{-2R}$.

V-4c) $\mathbf{E}\underline{U}\underline{U}^t = \mathbf{E}\underline{T}\underline{X}\underline{X}^t\mathbf{T}^t = \mathbf{T}\mathbf{R}\mathbf{T}^t$

V-4d) $\beta = \frac{1}{\sqrt{2}}$ pour que les lignes et colonnes soient normées. On trouve la matrice diagonale

$$\sigma_X^2 \begin{pmatrix} 1+\rho & 0 \\ 0 & 1-\rho \end{pmatrix}$$

V-4e) Comme la matrice d'autocorrélation de \underline{U} est diagonale, \mathbf{T} décorrèle la source. C'est la transformée (optimale) de Karhunen-Loève.

V-4f) $D^* = c\sigma_1\sigma_2 2^{-2R} = c\sigma_X^2 \sqrt{1-\rho^2} 2^{-2R}$

V-4g) Si $\rho = 0$, la source est blanche et la transformée n'améliore pas les performances. Quand $|\rho|$ augmente, la distorsion optimale D^* diminue. Le codage par transformée est donc d'autant plus efficace que la source est corrélée.



Cadre privé } sans modifications

Par le téléchargement ou la consultation de ce document, l'utilisateur accepte la licence d'utilisation qui y est attachée, telle que détaillée dans les dispositions suivantes, et s'engage à la respecter intégralement.

La licence confère à l'utilisateur un droit d'usage sur le document consulté ou téléchargé, totalement ou en partie, dans les conditions définies ci-après, et à l'exclusion de toute utilisation commerciale.

Le droit d'usage défini par la licence est limité à un usage dans un cadre exclusivement privé. Ce droit comprend :

- le droit de reproduire le document pour stockage aux fins de représentation sur un terminal informatique unique,
- le droit de reproduire le document en un exemplaire, pour copie de sauvegarde ou impression papier.

Aucune modification du document dans son contenu, sa forme ou sa présentation, ni aucune redistribution en tout ou partie, sous quelque forme et support que ce soit et notamment par mise en réseau, ne sont autorisées.

Les mentions relatives à la source du document et/ou à son auteur doivent être conservées dans leur intégralité.

Le droit d'usage défini par la licence est personnel, non exclusif et non transmissible. Tout autre usage que ceux prévus par la licence est soumis à autorisation préalable et expresse de l'auteur : sitepedago@enst.fr